

中文



STATE INTELLECTUAL PROPERTY OFFICE OF THE P.R.C

[HOME](#)[ABOUT SIPO](#)[NEWS](#)[LAW & POLICY](#)[SPECIAL TOPIC](#)[CHINA IP NEWS](#)**Title: Method and apparatus for data synchronization**

Application Number	01816387	Application Date	2001.09.21
Publication Number	1466723	Publication Date	2004.01.07
Priority Information	US09/677,3082000/9/29		
International Classification	G06F17/30		
Applicant(s) Name	Motorola Inc.		
Address			
Inventor(s) Name	A. Briancon		
Patent Agency Code	11219	Patent Agent	xie lina zhang lianshu

Abstract

A method and apparatus for providing data synchronization and/or resolving data conflicts is disclosed. In today's data communication environment multiple pieces of information may be received by a device and one of the pieces must be selected for processing or storage. The present system assigns a context value (135) to each piece of information (131) to describe the relevance of the piece of information (131) to a recipient. When a recipient recognizes that it has received two or more pieces of information (131) purporting to serve the same purpose, the context values (135) of the pieces of information (131) are evaluated and the piece of information (131) having the largest context value (135), representing the most relevance to the recipient, is selected.

[Machine Translation](#)[Close](#)



[12] 发明专利说明书

[21] ZL 专利号 01816387.4

[45] 授权公告日 2005 年 4 月 13 日

[11] 授权公告号 CN 1197028C

[22] 申请日 2001.9.21 [21] 申请号 01816387.4

[30] 优先权

[32] 2000. 9.29 [33] US [31] 09/677,308

[86] 国际申请 PCT/US2001/029695 2001.9.21

[87] 国际公布 WO2002/029640 英 2002.4.11

[85] 进入国家阶段日期 2003.3.27

[71] 专利权人 摩托罗拉公司

地址 美国伊利诺斯州

[72] 发明人 阿兰·布里昂松

审查员 盖 浩

[74] 专利代理机构 中原信达知识产权代理有限公司

代理人 谢丽娜 张天舒

权利要求书 2 页 说明书 9 页 附图 1 页

[54] 发明名称 数据同步方法及其装置

[57] 摘要

本发明公开了一种提供数据同步和/或解决数据冲突的方法和装置。在当今的数据通信环境中,设备可能接收到多条信息,并且必须从中选出一条信息以进行处理或存储。本发明的系统给每一个信息单元(131)分配一内容值(135)以描述信息(131)对接收者的相关性。当接收者识别出其接收到生成完成相同目的的两条或多条信息(131)时,计算信息(131)的内容值(135),并且选择具有最大内容值(135)、表示与接收者最相关的信息(131)。



I S N 1 0 0 8 - 4 2 7 4

1. 一种在用户信息响应设备中的信息同步的方法，该方法特征在于：

将内容字段分配给由用户设备访问的信息单元；

在所述用户设备处识别包括冲突信息的第一和第二信息单元；

分析第一和第二信息单元的内容字段以确定对第一和第二信息单元的各自用户的各自相关性；和

根据在分析步骤中确定的对所述用户的相关性来选择第一和第二信息单元中之一。

2. 根据权利要求 1 所述的方法，其中，分配给信息单元的所述内容字段包括表示信息单元所希望的用户数目的值。

3. 根据权利要求 2 所述的方法，其中，分配给信息单元的所述内容字段包括表示信息单元所希望的地理或逻辑范围的值。

4. 根据权利要求 2 所述的方法，其中，分配给信息单元的所述内容字段包括表示从用户方看的源的值。

5. 根据权利要求 4 所述的方法，其中，表征所述源的值表示所述源相对所述用户的关系。

6. 根据权利要求 1 所述的方法，其特征在干：

从源接收第一信息单元，并且从所述源接收分配给第一信息单元的内容字段。

7. 根据权利要求 6 所述的方法，其特征在干：接收分配了所述内容字段的信息单元之后，所述用户设备接收内容字段。

8. 根据权利要求 7 所述的方法，其特征在于：在接收分配了所述内容字段的信息单元之前，所述用户设备接收内容字段。

9. 一种由用户通过网络访问来自信息源的信息单元的方法，其特征在于：

在信息源处将内容字段分配给信息单元，所述内容字段表示所述信息单元对用户设备的可能相关性；

由所述信息源接收对信息单元和内容字段的请求；

分析来自所述用户设备的内容字段和分配给所述信息单元的内容字段；和

仅当所述分析步骤指出所述信息单元比从所述用户设备接收的所述内容字段示出的相关性与所述用户设备更加相关时，由所述信息源发送所述请求的信息单元到所述用户设备。

10. 根据权利要求 9 所述的方法，其中所述接收的内容字段表示对另一个信息单元的用户设备的相关性，所述另一个信息单元对于所述用户设备是已知的。

数据同步方法及其装置

5 技术领域

本发明涉及数据同步，具体地说涉及在要完成相同需要的多个信息单元中进行选择的方法和装置。

背景技术

10 现代社会中的数据量的产生、交换和使用产生了一些独特的问题。对于现代计算机网络来说，具有完成相同需要的多份数据是很常见的。例如，当用户下载文件到他或/她的掌上电脑（palm top）并且修改文件时，仅仅将修改的文件返回到下在原始文件的个人计算机（PC）。选择要跟踪（keep）的文件就是数据同步问题。

15 对于数据同步问题，最常用的方案是给所有的文件盖上日期戳，并且选择最近的文件用于后来的存储和使用。在 PC/掌上型电脑的例子中，在掌上型电脑或 PC 上的最近的版本是要保存的版本。当数据源的宽度增加时，例如连接到计算机网络如互联网，这种问题变得更加复杂。要求同步的数据可能来自很多源，且被全部打上日期戳。然而，接收者不知道哪一个接收数据的单元是可信的。由于数据是不是由其本身最近接收或发射的，数据比老的数据更相关，但是由不同源确认的好数据或准备的数据。

25 考虑不在办公室的销售人员需要知道待售产品的价格的情形。数据可从很多源接收到，所有的数据形成列表，从中销售人员可产生费用价格表。一个表可来源于营业处，另一个来源于市场部付总经理，第三份来源于销售人员的及时监管。这些不同源可以提供比准备或发射数据的确切时间更加相关的信息。在另一个例子中，数据不是来源于不同地方，但是可以发布往不同地方。一些数据可以寻址到大的地

30

理区域或者大量用户，同时潜在冲突的数据可以发往少数接收者或者可能只发往一个接收者。具有完成相同需要的数据（冲突数据）的接收者遇到很多不同的数据份数，从这些数据份数中不能利用当今的同步技术来适当地进行选择。因此，对于当今的数据选择方案，数据同步存在问

发明内容

根据本发明的一个方面，提供了一种在用户信息响应设备中的信息同步的方法。内容字段被分配给由用户设备访问的信息单元。接下来，在所述用户设备处识别包括冲突信息的第一和第二信息单元。然后，分析第一和第二信息单元的内容字段以确定对第一和第二信息单元的各自用户的各自相关性。之后，根据在分析步骤中确定的对所述用户的相关性来选择第一和第二信息单元中之一。

根据本发明的另一个方面，提供了一种由用户通过网络访问来自信息源的信息单元的方法。在信息源处将内容字段分配给信息单元，以表示所述信息单元对用户设备的可能相关性。接下来，由所述信息源接收对信息单元和内容字段的请求。然后，分析来自所述用户设备的内容字段和分配给所述信息单元的内容字段。之后，仅当所述分析步骤指出所述信息单元比从所述用户设备接收的所述内容字段示出的相关性与所述用户设备更加相关时，由所述信息源发送所述请求的信息单元到所述用户设备。

附图说明

- 图 1 是数据分布系统的方框图；
图 2 是数字网络消息的表示；
图 3 是用户和数据源之间的消息的消息序列图；
图 4 是示出源和用户之间的分布节点的数据分布系统的方框图；
图 5 是示出两个设备之间的同步消息消息序列图。

具体实施方式

根据本发明，解决了现有技术的问题，并且得到了提高。给信息单元分配一个内容值，该内容值反映了信息单元和分配给信息单元用户的可能相关性。当检测到冲突信息单元时，将它们的内容值进行比较，并且选择具有指出更大可能相关性的一个以用于处理和/或存储。

内容值可以由多个子内容值组成，子内容值在此指概念 (Concept)、作用域 (Scope)、范围 (Range)、焦点 (Focus) 和 Lobby。Concept 指信息单元要指向的用户数目。想要的用户数目越大，Concept 对内容的贡献越小。也就是说，最大的内容值来源于用于单个用户的信息单元。Range 是其分配给信息单元的地理或逻辑传播 (dispersion)。散布在全美国的信息单元具有的内容比仅仅在单个社区的单个 LAN 传播的信息单元的小。

Focus 限定了来自用户的信息单元的源。例如，当信息单元的源是老板时，Focus 值将导致相对高的内容值，而用于在用户公司外面的用户的 Focus 值则导致相对低的内容。

利用分配的内容来同步数据的方法可以被用来同步膝上型电脑和 PC，并且它可以用来在接收的信息单元之间选择哪一个满足用户需要。可以用信息单元来发射内容值到一个单元，该单元在此类信息单元之前或之后发射内容值。

在此介绍的实施例是用于由用户选择、处理和存储多个冲突的信息中之一的方法和装置。信息可以是任意的大小，但是希望本发明实施例采用具有例如大于 1 千字节的很大的最有用信息。冲突信息是用于解决相同的用户需要的信息。销售产品的当前价格表可以是一个很可能的冲突信息的例子。表示解决问题的多个可能的方案（或者互联网关键字搜索结果）的数据可以组成另一个例子。

图 1 表示一种环境，在此环境中，从包括存储器 103 的用户计算机 101 来看可能出现信息冲突。用户计算机 101 连接到掌上型设备 105，其可将信息提供给用户。来自掌上型设备 105 的信息可以是原
5 是信息或者是之前从用户计算机 101 的存储器 103 下载下来的、被修改并返回给用户计算机的信息。在返回时，可能出现关于版本的冲突，即存储器版本或掌上型设备版本被存储在存储器 103 和/或由用户计算机 101 进行处理。

用户计算机 101 也经网络 107 连接到多个信息源，其中特别地示
10 出了信息源 109 和 111。例如，网络 107 可以是公知的互联网，且源 109 和 111 可以是连接到互联网的服务器。源 109 和 111 可以分开地分布得很广，并且可以由不同的人们和小组进行控制。当用户计算机 101 识别出其接收到两条信息或识别出两条信息完成相同的用户需要时，在用户计算机处出现信息冲突。

15 在本发明的实施例中，内容字段与各条信息相关。用户计算机 101 确定跟踪或处理哪一条信息，根据内容字段来确定丢弃哪一条信息。该内容字段向用户计算机指出相关的信息到用户的可能相关性。当第一信息的可能的用户相关性（内容）大于第二冲突的信息的可能的用户相关性时，将保持和/或处理第一信息，并且丢弃第二信息。
20

图 2 表示当发射这些信息时，与信息相关的字段分配。图 2 包括
25 标记字段 121、数据字段 123 和内容字段 125。数据字段 123 表示与标记和内容字段相关的信息。标记字段 121 包括源和目的地址以及关于数据字段 123 的、定义的变量的其它通信。内容字段 125 包括可从
中确定对用户的内容或相关性的信息。图 2 是以串行的字符串表示的
标记、数据和内容字段，其可能是通过网络传送信息的方式。在一个
实施例中，内容字段可以存储在具有信息的源和目的地址上的查找表
30 中。在另一个实施例中，可以如下面介绍的一样，从其它消息属性中
导出内容字段 125 或者其中的部分。

在本发明中，信息内容包括与相关的信息的可能相关性相关的单调上升函数 (monotonic up functions)。对于给定的传送信息，内容具有一值，可将该值与冲突信息的内容值比较以识别跟踪或处理哪一信息。在下面的例子中，选择具有最大内容的信息。需要提到的是，内容具有意义和相对的感知 (sense)。也就是，当与另一个内容值或预定的参考内容值比较时，内容值变得重要 (significant)。

内容 5 由一个或多个不同值组成，下面介绍这样的例子。下面的公式表示考虑了 5 个值的内容定义：

Context = monotonic up function (
monotonic up function (-concept),
monotonic up function (-focus),
monotonic up function (scope),
monotonic up function (lobby)).

Concept 表示希望通信的用户数目，concept 值越大，则接收信息的用户越多。由于在内容定义中 concept 是负函数，希望通信的用户数目越少，对内容的贡献就越大。也就是，当向单个用户传送通信时，内容将大于当向 10 个用户传送通信时的内容。

在通信中，concept 可以是隐式的或者显示的。显示的 concept 是从信息源用通信发射出来的一个 concept。这样的显示 concept 可以作为部分内容与通信一起发送，或者可以在来自接收者的请求上提供。图 3 表示在源 (例如 111) 和用户 (例如 101) 之间传递消息以请求一 concept 值。最初，源将信息 131 发射到用户 101。用户识别可能的冲突，并且识别合乎要求来估计接收的信息的 concept 值。在这种情况下，用户将 concept 请求消息 133 返回到源 111，源 111 通过将 concept 值 125 发射到用户来进行响应。然后，用户 101 将接收的 concept 值 135 和冲突信息相关，并且适当地选择冲突信息的版本以供使用。

构成内容字段的另一个值是 **Range** (范围), 其是信息的地理或者逻辑分布。分布的区域越大, 则 **range** 越大, 且由于 **range** 是内容的负函数, 则 **range** 越大, 内容越小。通常而言, 分布很广的信息

5 (something) 相比于分布很窄的信息很可能具有对各个用户较小的相关性。

根据基本的范围单元 (**range unit**) 来给 **range** 分配一个值, 在本发明的实施例中基本 **range** 单元是单个人计算机 (PC)。下表表示用于各个 **range** 的范围值:

10

<u>范围单元</u>	<u>范围值</u>
个人计算机 (PC)	1
LAN 段	段上的 PC 号
LAN	LAN 上的 PC 号
15 蓝牙 Pod	n, n 根据数据范围变化
广域网	N, 网络覆盖的平方 公里 (最低值 N) 最高值 n)

由源进行前向估计可以提供 **range**, 其跟踪传送节点的数目, 并且将每一个节点乘以从其类型确定的权重因子。图 4 表示具有节点的网络, 为这些节点确定前向估计。在每一个分布树或者沿从源到用户的路径的节点处、根据分支数目来产生估计。因此, 如图所示, 当开始分支数目是 $N1$, 第二分支数目是 $N2$ 和最后的接收者数目是 NP , 则 **range** 值是 $N1 \times N2 \times NP$ 。

20

25

也可以进行 **range** 的反向估计。对于反向估计, 分布 (存储和发送) 链的每一个元件 (element) 将它的通信 **range** 报告给源。在链中的每一个元件将其所有输入按降序排列, 并且将它们相加起来。源 (例如如图 4 的 111) 从元件接收所有的“报告返回”, 且计算随后发送到接收用户的 **range** 值。与 **Concept** 一样, **range** 可以与要发送的信息一起

30

发送，或者在之后发送。

Focus 值限定从用户方看的源。可以按照下表定义 focus:

	<u>到用户的源</u>	<u>焦点</u>
5	老板	1
	工作对等 (peer)	n
	同事	m
	局外人	p

10 其中, $1 < n < m < p$ 。

由于 focus 是负的单调函数, 具有 1 的老板 focus 贡献用于最大的内容值, 而具有 p 的最大外部 focus 贡献用于最小的内容值。从上表可见, 按照控制和/或工作关系来说, 源离用户越远, 则 focus 对内容值的贡献越小。Foccus 也可以是由源或用户计算的隐式函数。例如, 15 由于 focus 与源到用户的工作选择相关, focus 可以从跳跃 (hop) 数目中或者防火墙 (firewall) 中得出, 通过防火墙将信息从源传到用户。

Scope 被定义为要传送的信息的内容定量说明。信息的集中度越窄, 则 Scope 越大, 相反也成立。Scope 可以被预先确定并且存储在内容源 (服务器) 处的表中, 并且可以被访问以与内容一起发送。当它被发送时, 也可以由数据的自然语言处理来以接近实时地设置 Scope。

25 Lobby 是由源任意地设置的值, 以反映从源看到的与发送的数据的可能相关性。因此, 对于很大的相关性, 源可为数据分配大的 Lobby 值, 并且可以给潜在的较小相关数据分配相对低的值。为了防止源填满其自身的内容值而排除输入其它的内容值, 必须给分配给 Lobby 的值加上限制, 这样的限制不能用来拒绝其它值。

30

从上面所述可知，可以使用与多条信息相关的内容值来决定存储和/或处理哪一信息。在奥马哈分公司的销售人员可能需要尺寸为从 A 到 ZZZ 的小配件（widget）的当前价格表，并且将在他或她的计算机中的信息代表它们。销售人员可能接收到声明要更新价格表的第一条和第二条信息。第一条信息可能来自位于纽约本部的市场部，这个市场部将该信息发布到分布在全国的、包括奥马哈分公司的 100 个分公司。第一价格信息具有由大的 Concept（很多接收者）、很宽的 range（全国和/或很多防火墙）、相对大的 focus（来自远方的同事）组成的相对低的内容。

10

第二条价格表信息可能来自在与该销售人员相同的 LAN 中的销售人员的老板、并且仅向 LAN 上的 6 个销售人员发送。第二条信息具有由大的 Concept（几个用户）、相对窄的 range（都在 LAN 上）和低的 focus（来自老板）组成的相对大的内容。因此，根据第一条和第二条信息的内容，选择第二条来代替或更新销售人员的当前价格表文件。

上面的例子涉及到从多个源接收数据，并且允许根据 Context 选择数据。内容字段的优点可以成功地应用于需要选择数据的其它场合。例如，当两个设备（图 5）存储前在冲突的数据，并且希望在两个设备上保存相同的数据时，设备可以共享内容值，并且可以用在具有最大内容的一个或其它设备上的数据来更新其它设备。另一个应用的例子是，用户已经具有访问互联网服务器的信息以识别是否需要更新在用户处的信息。这样的情形可发生在当浏览器访问已经存储在用户计算机上的网页时。可以将服务器上的网页的内容字段与用户的内容字段进行比较，仅当服务器上的内容大于用户上的内容时进行更新。作为附加的一个例子，考虑通过计算机网络进行分布搜索数据的情形，可以由互联网搜索引擎来执行搜索。由于搜索引擎识别有各种可能的候选文件，收集它们的内容值并且用来选择用于用户考虑的候选文件或者进行分等级。

30

尽管已经说明和介绍了本发明的特定实施例，本领域普通技术人员将认识到，可以进行很多的改变和修改，并且希望权利要求书覆盖落在本发明范围内的那些改变和修改。

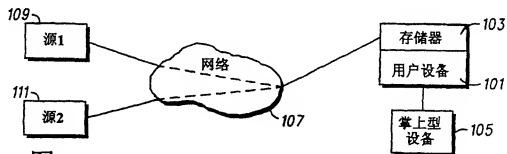


图1

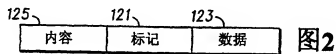


图2

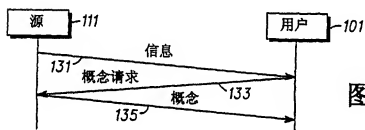


图3

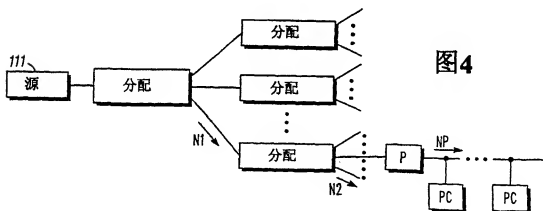


图4

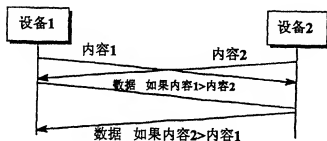


图5